1 *Burkholderia humptydooensis* sp. nov., a *Burkholderia thailandensis*-like species and the fifth

2 member of the *pseudomallei* complex.

3 Apichai Tuanyok[1, 7]
4 Mark Mayo[2]
5 Holger Scholz[3]
6 Carina Hall[1]
7 Christopher J. Allender[1]
8 Mirjam Kaestli[2]
9 Jennifer Ginther[1]
10 Senanu Spring-Pearson[1]
11 Molly C. Bollig[1]
12 Joshua K. Stone[1]
13 Erik W. Settles[1]
14 Joseph D. Busch[1]
15 Lindsay Sidak-Loftis[1]
16 Jason W. Sahl[1]
17 Astrid Thomas[3]
18 Lisa Kreutzer[3]
19 Enrico Georgi[3]
20 Jay E. Gee[4]
21 Richard A. Bowen[5]
22 Jason T. Ladner[6]
23 Sean Lovett[6]
24 Galina Koroleva[6]
25 Gustavo Palacios[6]
26 David M. Wagner[1]
27 Bart J. Currie[2]
28 Paul Keim[1*]
29
30 [1] Department of Biological Sciences and Center for Microbial Genetics and Genomics, Northern
31 Arizona University, Flagstaff, Arizona 86011, USA.
32
33 [2] Menzies School of Health Research, Charles Darwin University, and Northern Territory
34 Medical Program, Royal Darwin Hospital, Darwin, Northern Territory, Australia.
35
36 [3] Bundeswehr Institute of Microbiology, Neuherbergstrasse 11, D-80937 Munich, Germany
37
38 [4] Bacterial Special Pathogens Branch, Division of High-Consequence Pathogens and Pathology,
39 Centers for Disease Control and Prevention, 1600 Clifton Rd., Atlanta, GA 30333, USA.
40
41 [5] Department of Biomedical Sciences, Colorado State University, Fort Collins, Colorado 80521,
42 USA.
43
44 [6] Center for Genome Sciences, USAMRIID, Fort Detrick, MD, USA
45

46 [7]Current address: Department of Infectious Diseases and Pathology, College of Veterinary
47 Medicine, University of Florida, Gainesville, FL 32608
48
49 *Corresponding author: Email: Paul.Keim@nau.edu, Center for Microbial Genetics and
50 Genomics, Northern Arizona University, Flagstaff, Arizona 86011, USA.
51
52 **Running title:** *Burkholderia humptydooensis* sp. nov.

53 **Abbreviations:** BCC, MLST, FAME, MALDI-TOF, GGDC, ASA

54 **Footnote:** GenBank accession numbers for the 16S rRNA gene sequences of *B. humptydooensis*

55 sp. nov. strains MSMB121, MSMB122, are KF378608 – KF378609, respectively. The complete

56 whole genome sequence of the strain MSMB121 was published under GenBank accession nos.

57 CP004095 and CP004096.

58 MSMB43: GCA_001513745

59 MSMB122: SRA (SRR1956040); LOYF00000000

60

61 **Summary**

62 During routine screening for endemic *Burkholderia pseudomallei* from water wells in northern

63 Australia, Gram-negative bacteria (strains MSMB43^T, MSMB121, and MSMB122) with a

64 similar morphology and biochemical pattern to *B. pseudomallei* and *B. thailandensis* were co-

65 isolated with *B. pseudomallei* on Ashdown's selective agar. To determine the exact taxonomic

66 position of these strains and to distinguish them from *B. pseudomallei* and *B. thailandensis*, they

67 were subjected to a series of phenotypic and molecular analyses. Biochemical and fatty acid

68 methyl esters analysis was unable to distinguish *B. humptydooensis* sp. nov. from closely related

69 species. In MALDI-TOF analysis, all isolates grouped together in a cluster separate from other

70 *Burkholderia* spp. 16S rRNA and *recA* sequence analysis demonstrated phylogenetic placement

71 for *B. humptydooensis* sp. nov. in a novel clade within the *B. pseudomallei* group. MLST

72 analysis of the three isolates in comparison with MLST data from 3,340 *B. pseudomallei* strains

73 and related taxa revealed a new sequence type (ST318). Genome to genome distance calculations

74 and average nucleotide identity of all isolates to both *B. thailandensis* and *B. pseudomallei*,

75 based on whole genome sequences, also confirmed *B. humptydooensis* sp. nov. as a novel

76 *Burkholderia* species within the *pseudomallei* complex. Molecular analyses clearly demonstrate

77 that strains MSMB43^T, MSMB121, and MSMB122 belong to a novel *Burkholderia* species for

78 which the name *Burkholderia humptydooensis* sp. nov. is proposed with the type strain

79 MSMB43^T (=ATCCXXXXX^T; LMGXXXX, CP013380-CP013382).

80   *Burkholderia* species are abundant and occupy diverse ecological niches, including soil,

81 plants, animals, and humans. Probably the most diverse and environmentally adaptable plant-

82 associated bacteria also belong to the genus *Burkholderia* (Compant et al., 2008). Many species

83 of *Burkholderia* have been described since the discovery of *B. cepacia* by W. H. Burkholder in

84   1949 as the cause of onion rot (Burkholder, 1950), and later recognized as a human pathogen.

85   Currently, there are more than 90 identified species in this genus (Vial et al., 2011, Gu et al.,

86   2015). There has been a proposal to divide the species into two genera, one of which would

87   retain the *Burkholderia* name and the other *Paraburkholderia* gen. nov. (Sawana et al., 2014). At

88   least 17 closely-related species belong to the *Burkholderia cepacia* complex (BCC), including

89   nine well-known species comprising the genomovars I-IX, respectively: *B. cepacia*, *B.*

90   *multivorans*, *B. cenocepacia*, *B. stabilis*, *B. vietnamiensis*, *B. dolosa*, *B. ambifaria*, *B. anthina*, *B.*

91   *pyrrocinia* (Coenye and Vandamme, 2003); seven recently identified species: *B. latens*, *B.*

92   *diffusa*, *B. arboris*, *B. seminalis*, *B. metallica*, *B. contaminans*, and *B. lata* (Vanlaere et al., 2009,

93   Vanlaere et al., 2008); and *B. ubonensis* (Yabuuchi et al., 2000). Most of these species reside in

94   soils and are associated with plants.

95   Notably, there are two *Burkholderia* species that can cause severe human and animal

96   diseases: *B. pseudomallei* and *B. mallei*, the causative agents of melioidosis and glanders,

97   respectively. *B. pseudomallei* is a major cause of community acquired sepsis in Northeast

98   Thailand and northern Australia (Cheng and Currie, 2005). Due to the concerns of their potential

99   use as weapons of mass destruction, federal health agencies in the United States have recently

100   classified these species as Tier 1 (Top Tier) disease agents (Butler, 2012). It has been well

101   established that *B. mallei* is a clone of *B. pseudomallei* that became a host-adapted pathogen in

102   equines resulting in massive genome reduction (Godoy et al., 2003). Genetically, both species

103   are members of the *pseudomallei* phylogenetic group or complex

104   (http://www.ncbi.nlm.nih.gov/Taxonomy/). Three additional closely related species have been

105   identified so far in this group, including *B. thailandensis* (Brett et al., 1997), *B. oklahomensis*

106   (Glass et al., 2006b), and a newly identified *B. thailandensis*-like species (Currie, 2015, Gee et

107 al., 2008). These closely-related species are soil saprophytes and are considered non-pathogenic,

108 although a few strains of *B. thailandensis* and *B. oklahomensis* have been described as causing

109 clinical infection in humans (Glass et al., 2006a, Glass et al., 2006b).

110

111 Strain MSMB43[T] (proposed type strain) was isolated from a water sample from an

112 automated water bore (well) collected in 1995 and examined for *B. pseudomallei* in the Northern

113 Territory (NT) of Australia (Gee et al., 2008). It was initially thought to be *B. thailandensis*

114 because this strain was able to assimilate arabinose as a sole carbon source, which is a trait used

115 to discriminate *B. thailandensis* from *B. pseudomallei* (arabinose negative) (Gee et al., 2008).

116 The bore from which MSMB43[T] was discovered is located in Humpty Doo, a region of rural

117 properties outside the capital of the NT, Darwin. The Top End of the NT has a high incidence

118 rate of melioidosis (Parameswaran et al., 2012). Indeed the water sample from which MSMB43[T]

119 was recovered also yielded *B. pseudomallei.* An additional two strains (MSMB121 and

120 MSMB122) were both isolated in 2007, from a single separate bore water sample within the NT

121 collected approximately 950 km south of the territory capital, Darwin, resulting in a 910 km

122 separation between the two sample sites of MSMB121/MSMB122 andMSMB43[T]. To date, *B.*

123 *humptydooensis* sp. nov. has not been identified outside the NT and it has not been isolated from

124 any clinical specimens from patients within the NT.

125 The specific epithet "*humptydooensis*" given to this new species was adopted from the

126 location name "Humpty Doo," the location where this new species was first discovered. The

127 name "humptydooensis" has been used previously for a new species of beetle (*Australoxenella*

128 *humptydooensis)* from this same geographic region (Howden and Storey, 1992).

129

130   As described previously (Gee et al., 2008), MSMB43[T] did not grow while incubated at a

131   temperature greater than 42°C and also produced little or no gas from nitrate. All three strains

132   were grown at 8, 25, 37, 42, and 45°C for 24, 48, 72, and 144 h on Columbia blood agar,

133   MacConkey agar, Ashdown's selective agar, and Standard I with and without supplementary

134   CO_2. On Columbia blood agar smooth and creamy-white colonies were observed after 24 h,

135   whereas red, convex and small (1-2mm) colonies were observed on MacConkey after 48h. Dry

136   and wrinkled colonies were observed on Ashdown's agar after 72 h of growth (Fig. 1), similar to

137   the appearance of *B. pseudomallei*, while slimy-confluent honey-like growth appeared on

138   Standard I after 48h (Fig. 1). Bacterial growth was visible on all media after incubation at 25-

139   42°C for at least 24 h with best growth observed on Columbia blood agar. No growth was

140   observed at 8°C and 45°C. The optimal temperatures for growth were between 28-37°C

141   aerobically.

142   Cell morphology was examined using a Zeiss light microscope at 1000x magnification

143   with cells grown for 2 days at 37 °C. All strains showed a Gram-negative bipolar staining

144   appearing as rods of 2-3 µm in length and 0.4-0.8 µm in diameter. All strains were motile in

145   semi-solid media.

146   Biochemical data were obtained for all three strains of *B. humptydooensis* sp. nov.

147   (MSMB43[T], MSMB121, and MSMB122) and compared to the type strains of *B. pseudomallei*

148   (K96243) (Wuthiekanun et al., 1996) and *B. thailandensis* (E264) using the API NE and API

149   Zym (bioMérieux) systems according to the manufacturer's instructions. Differentiation of *B.*

150   *humptydooensis* sp. nov from *B. thailandensis* was not possible based on the biochemical

151   reactions (Table 1).

152   Matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-

153   TOF MS) was performed on all three *B. humptydooensis* sp. nov strains and they clustered with

154   other members of the *B. pseudomallei* complex (see SI doc and Fig. S1). Fatty acid methyl esters

155   analysis was unable to distinguish among the fatty acid profiles from the three *B.*

156   *humptydooensis* sp. nov. strains and its closely related species (five *B. thailandensis*, two *B.*

157   *oklahomensis*, and three *B. ubonensis* strains) (see SI doc and Fig. S2).

158   Minimum inhibitory concentrations (MICs) were determined by broth microdilution

159   method using commercially available CE-certified MICRONAUT-S 96-well microtitre plates

160   (Merlin, Bornheim-Hersel, Germany) containing a 2-fold serial dilution of the following

161   antibiotics: amoxicillin/clavulanic acid (0.5–64/0.25–32 mg/L), ceftazidime (0.5–64 mg/L),

162   imipenem (0.25–32 mg/L), rifampicin (0.0625–8 mg/L), chloramphenicol (0.5–64 mg/L),

163   trimethoprim/sulfamethoxazole (0.25–32/4.75–608 mg/L), streptomycin (0.25–32 mg/L),

164   gentamicin (0.25–32 mg/L), doxycycline (0.25–32 mg/L), tigecycline (0.03125–4 mg/L),

165   ciprofloxacin (0.03125–4 mg/L), and levofloxacin (0.0625–4 mg/L). One well without antibiotic

166   was used as a growth control. All plates containing the lyophilized antimicrobial substances were

167   stored at room temperature until use.

168   Testing conditions were in accordance with the current Clinical & Laboratory Standards

169   Institute (CLSI) recommendations for *B. pseudomallei* (CLSI M45-A2 Volume 30 No. 18).

170   Single colonies of MSMB43^T, MSMB121, and MSMB122 were picked from agar plates and

171   inoculated in physiological saline until the turbidity matched that of a 0.5 McFarland standard.

172   The suspension was diluted 221-fold in cation-adjusted Mueller Hinton II broth (Becton

173   Dickinson, catalog #297701). After incubation for 24 h at 37°C in a 5% CO_2 atmosphere

174   bacterial growth was verified photometrically at a wavelength of 620 nm using a commercial

175 photometer (MERLIN, Bornheim-Hersel, Germany) and each strain was tested in triplicate.

176 Additionally, a gradient strip method (Etest, bioMérieux) was applied to investigate a broader

177 range of antibiotic concentrations.

178 Adopting the CLSI breakpoints of *B. pseudomallei*, all strains were determined to be *in*

179 *vitro* susceptible to ceftazidime, imipenem, trimethoprim/sulfamethoxazole, and doxycycline,

180 whereas a resistance to amoxicillin/clavulanic acid was observed. (Table 2). The antimicrobial

181 susceptibility pattern of *B. humptydooensis* sp. nov. generally resembles that of *B. pseudomallei*

182 (Ahmad et al., 2013, Thibault et al., 2004, Trunck et al., 2009, Crowe et al., 2014), except for the

183 elevated MICs of amoxicillin/clavulanic acid (Table 2). No significant differences were observed

184 either among the three strains or between the two different susceptibility testing methods. As the

185 maximum concentration of aminoglycosides in the microtiter plates was 32 mg/L, high-level

186 streptomycin resistance but low-level gentamicin resistance was confirmed using the Etest

187 method (data not shown).

188 The pathogenic potential of *B. humptydooensis* sp. nov., MSMB43^T, was investigated in a

189 BALB/c mouse model and compared to *B. thailandensis* (type strain E264). Live culture was

190 grown to logarithmic phase (OD$_{600}$ ~ 1.0) in Luria-Bertani (LB) broth as described in (Morici et

191 al., 2010). Sterile 1xPBS was used to wash cells twice before making dilutions for injecting

192 mice. Viability counts of the final inocula were made on LB agar plates. Six to eight week old

193 female BALB/c mice in treatment groups of 5 mice per cage were used. Food and water were

194 provided *ad libitum*. All mice in a single cage received the same infectious dose (*B.*

195 *humptydooensis* sp. nov.: 1.05 x 10^4, 10^5 or 10^6 CFU; *B. thailandensis*: 3.4x 10^4, 10^5 or 10^6

196 CFU) via a single subcutaneous (s.c.) injection in the scruff of the neck. Mice were monitored

197 daily for health status. All mice were euthanized on day 21 post injection. This work was

198   conducted under approved protocols from the NAU IACUC (Protocol 14-011) and DOD

199   ACURO (HDTRA1-12-C-0066_Wagner). Neither species caused mortality in any mice when

200   delivered via the s.c. route, nor did any mice show outward signs of illness. In comparison, s.c.

201   infections of fully virulent *B. pseudomallei* results in 50% mortality within 10 days at a dose of

202   10^3 CFU (Barnes and Ketheesan, 2005). It remains unknown if the inhalation route increases the

203   pathogenicity of species tested in the same way as *B. thailandensis* E264, which can cause high

204   mortality in mice at doses of 10^4 - 10^6 CFU when delivered as an aerosol (Morici et al., 2010,

205   West et al., 2008, Wiersinga et al., 2008).

206   16S rRNA and *recA* gene sequencing analyses were performed on three *B.*

207   *humptydooensis* sp. nov. strains: MSMB43^T, MSMB121, and MSMB122 as previously described

208   (Brett et al., 1997, Payne et al., 2005). From whole genome analysis of strain MSMB43^T using

209   SSU-ALIGN (Nawrocki et al., 2009) we determined that four rRNA operons are present on its

210   chromosomes of which two unique versions were found (AQ610_12930, AQ610_01425;

211   AQ610_21350, AQ610_02540). These two copies of the 16S rRNA genes are different, which

212   leads to ambiguities in conventional sequencing (see Fig. S3). The 16S rRNA gene sequence

213   similarities of *B. humptydooensis* sp. nov. to other members of the *pseudomallei* complex (*B.*

214   *thailandensis*, *B. mallei*, *B. oklahomensis*) were 99%. Phylogenetic reconstruction of 16S rRNA

215   and *recA* sequences using MEGA version 6 (Tamura et al., 2013) confirmed genetic proximity to

216   the *pseudomallei* complex but also determined that all *B. humptydooensis* sp. nov. strains formed

217   their own group within this complex (see Fig. S3-S4).

218   MLST was performed on all three *B. humptydooensis* sp. nov. strains as previously

219   described (Godoy et al., 2003). As of October 6, 2015, a total of 1,429 sequence types (STs) had

220   been identified in *B. pseudomallei* and closely related species by MLST (http://www.MLST.net).

221 All three *B. humptydooensis* sp. nov. strains, MSMB43[T], MSMB121, and MSMB122 are ST318

222 and there are no other representatives of this ST. Overall, phylogenetic analysis using MLST

223 data supports the separation of *B. humptydooensis* sp. nov. from the other *pseudomallei* complex

224 members, as described previously (Gee et al., 2008).

225   Two genomes were sequenced on the PacBio platform, resulting in one finished assembly

226 (MSMB43[T]) and one mostly finished assembly with 4 contigs (MSMB122) (Table 3). Two other

227 genomes that group with *B. humptydooensis* sp. nov. are present in GenBank, see Figure 2 (*B.*

228 *thailandensis* MSMB121and *Burkholderia* spp. MSMB43 (2002721687) with GenBank

229 assembly accession numbers of GCA_000385525 and GCA_000959325 respectively). The

230 MSMB43[T] genome had one circular contig ~305Kb long that appears to be a plasmid; this same

231 sequence is also present in the completed genome of *Burkholderia* spp. MSMB43 (2002721687).

232 A comparative genomics approach using LS-BSR (Sahl et al., 2014) demonstrated that a large

233 stretch of the *B. pseudomallei* K96243 genome (BPSS1165 – BPSS1184) on chromosome 2 is

234 highly conserved (>98% identity) in the plasmid sequence, suggesting a shared origin for these

235 regions.

236   For the core genome phylogeny, genomes were aligned against *B. pseudomallei* K96243 with

237 NUCmer (Delcher et al., 2003). The reference genome was also aligned against itself to identify

238 duplicated regions, which were masked from subsequent analyses; the NASP pipeline was used

239 to wrap these methods (http://tgennorth.github.io/NASP/). A phylogeny was inferred with

240 RAxML v8 (Stamatakis, 2014) on a large set (n=331,000) of concatenated SNPs using a time

241 reversible model incorporating the Lewis ascertainment bias correction. The phylogeny

242 demonstrates the position of *B. humptydooensis* sp. nov. in relation to other clades in the *B.*

243 *pseudomallei* complex and confirms the results from other methods (Fig. 2).

244

245   DNA-DNA hybridization (DDH) is the current gold-standard for bacterial species

246   delineation. DDH is necessary for the description of a new species within a taxon when strains

247   share more than 97% of 16S rRNA gene sequence similarity (Goebel and Stackebrandt, 1994,

248   Tindall et al., 2010). If DNA-DNA relatedness between two strains is less than 70%, both strains

249   are considered as different species. However, DDH is laborious and difficult to standardize and

250   inter-laboratory reproducibility is relatively low. DDH was previously performed on MSMB43[T]

251   and has a relative binding ratio of 91% with *B. thailandensis* (ATCC 700388T) with a

252   divergence of 4% (Gee et al., 2008). Because of the drawbacks of conventional DDH and the

253   rapid progress in genome sequencing techniques, various *in silico* algorithms for calculating

254   genome-to-genome similarities or distances have been developed. In addition to the commonly

255   used average nucleotide identity (ANI) (Konstantinidis and Tiedje, 2005), recently a highly

256   reliable estimator for the relatedness of genomes was developed by Jan P. Meier-Kolthoff and

257   colleagues (Meier-Kolthoff et al., 2013). Genome to genome distance calculation (GGDC)

258   produces digital DDH values that correlate well with values obtained by conventional DDH,

259   which is of utmost importance for compatibility with the current species concept, and also

260   provides confidence-interval estimation. Values reported in Table 4 were generated using

261   formula 2, which summed the identities found in high-scoring segment pairs (HSP) and then

262   were divided by the overall HSP length. The GGDC service is available from the German

263   Collection of Microorganisms and Cell Cultures homepage (http://ggdc.dsmz.de/distcalc2.php).

264   PacBio assemblies were used to determine the distances among *B. humptydooensis* sp. nov. and

265   other closely related species, *B. humptydooensis* sp. nov. strains MSMB43[T], MSMB121, and

266   MSMB122. The genomes were subjected to GGDC analysis and compared to the available

267   genome sequences of *B. pseudomallei*, *B. oklahomensis* and *B. thailandensis* reference strains

268   (K96243, C6786^T, and E264, respectively). For comparison, the average nucleotide identity

269   (ANI) values were also calculated for all reference sequences with JSpecies (Richter and

270   Rossello-Mora, 2009); the authors of JSpecies determined that ANI values <95% indicate

271   separate species.

272   Among the three tested *B. humptydooensis* sp. nov. genomes, the calculated GGDC and

273   ANI values were in the range of 93-99% and 98-99%, respectively (Table 4), indicating that all

274   of these tested strains belong to a single species, including the proposed *B. humptydooensis* sp.

275   nov. type strain MSMB43^T. As expected from WGS analyses, strain MSMB43^T had a slightly

276   lower GGDC similarity (93%) to the other two *B. humptydooensis* strains, which were

277   approximately 97% similar. Further GGDC analysis (Table 4) has determined that the

278   similarities of all *B. humptydooensis* strains compared to all other tested *Burkholderia* species in

279   the *B. pseudomallei* complex was less than 70%, with the highest similarity detected to *B.*

280   *thailandensis* (51.1% (± 3.2%)). This confirms that the three tested strains are not *B.*

281   *thailandensis*, but rather a distinct species. The GGDC similarity between *B. mallei* and *B.*

282   *pseudomallei* was 92.5% which confirms previous conventional DDH results and demonstrates

283   that, from a strict taxonomic point of view, they belong to a single species.

284   **Description of *Burkholderia humptydooensis* sp. nov.**

285   *Burkholderia humptydooensis* sp. nov. (hump.ty.doo.en'sis. L. gen. adj. *humptydooensis*,

286   pertaining to Humpty Doo, a small town in Northern Territory of Australia, where the first

287   member of this species was isolated).

288   Bacilli, 0.4-0.8 µm in diameter and 2-3 µm in length, arranged individually or in irregular

289   clusters. The organism is Gram-negative with bipolar staining, motile, and non-spore forming.

290   Growth is observed in a temperature range of 25-42°C within 24-48 h on various standard solid

291   media. Within 24 h, small colonies (0.5-1 mm) are formed on non-selective media (Columbia

292   blood and Standard-I) and after 48 h also on selective media (Ashdown's, MaConkey). Best

293   growth occurs at 28-37°C after ≥ 24 h. Colonies become confluent and honey-like in appearance

294   on glycerol- containing medium (ST-I) after 48 h. On Ashdown's selective agar, highly wrinkled

295   purple colonies are observed within ≥ 48 h, resembling the growth of *B. pseudomallei*.

296   Positive assimilation (API 20NE) was found for D-glucose, L-arabinose, D-mannose, D-

297   mannitol, N-acetyl-glucosamine, potassium gluconate, capric acid, adipic acid, malic acid,

298   trisodium citrate, and phenylacetic acid whilenegative for D-maltose. Aesculin and gelatin are

299   hydrolyzed. Variable reactions are L-arginine and 4-Nitrophenyl-β D-galactopyranoside (PNPG).

300   Positive (API ZYM) for alkaline phosphatase, esterase, esterase lipase, lipase, leucine

301   arylamidase, acidic phosphatase, and naphthol-AS-BI-phosphohydrolase. Enzymes absent on

302   API ZYM are valine arylamidase, cystin arylamidase, trypsin, α-chymotrypsin, α- and β-

303   galactosidase, β-glucuronidase, α- and β-glucosidase, N-acetyl-β-glucosaminidase, α-

304   mannosidase, and α-frucosidase. This species is aerobic, catalase and oxidase-positive, urease

305   and indole negative. Nitrate and nitrite are reduced (with no gas formation from nitrite) and no

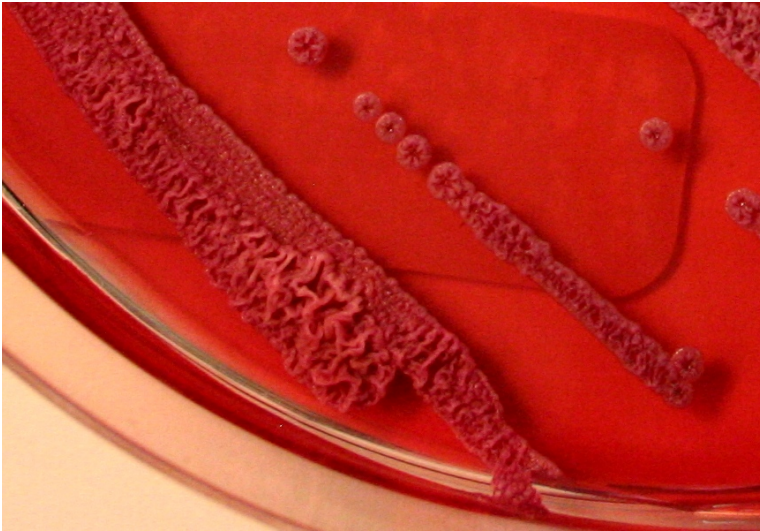306   production of H_2S.

307   *B. humptydooensis* sp. nov. strains are resistant to aminoglycosides and

308   amoxicillin/clavulanic acid, but susceptible to trimethoprim/sulfmethoxazole, doxycycline,

309   imipenem, and ceftazidime. All *B. humptydooensis* sp. nov. strains are seroreactive with sera

310   from melioidosis patients who were infected with *B. pseudomallei* serotype B strains. All strains

311   produced O-antigen ladder type B2, except that strain MSMB43[T] produced a novel O-antigen

312   ladder type (Stone et al., 2012). The type strain, MSMB43[T], has been previously referred to as *B.*

313   *thailandensis*-like species in multiple studies (Currie, 2015, Gee et al., 2008). MSMB43^T was

314   isolated in 1995 from an automated water well (bore) in Humpty Doo, Australia. Of interest, *B.*

315   *pseudomallei* was also cultured from that same water sample. *B. humptydooensis* sp. nov., like *B.*

316   *thailandensis*, is non-virulent in mice. In addition, MSMB43^T is known to produce

317   Thailanstatins, which possess antiproliferative activities in representative human cancer cell lines

318   (Liu et al., 2013). The type strain MSMB43^T has been deposited in the American Type Culture

319   Collection as ATCC XXXXX and the Belgian Co-ordinated Collections of Micro-organisms as

320   LMGXXXX.

326

327

328   **Figure 1.** Colony morphology of *B. humptydooensis* sp. nov. MSMB43^T.

329   (a) Culture grown on Ashdown's agar and on (b) Standard I.

330

331   (b)

332

333

334

335

336

337   **Figure 2.** Core genome phylogeny of *B. humptydooensis* sp. nov.

338   SNPs from the comparison of four *B. humptydooensis* sp. nov. genomes and representatives of
339   the other closely related species were used to reconstruct the phylogenetic relationships.
340   Genomes from this study in bold and assembly numbers in parentheses. Numbers at nodes
341   indicate bootstrap support values. Collapsed nodes shown in gray.

342

343

344

345 **Table 1.** Phenotypic characteristics of *B. humptydooensis* sp. nov. and closely related species
346 within the *B. pseudomallei* group.

347 Strains: Bp, *Burkholderia pseudomallei* (K96243)*; Bt, *B. thailandensis* (E264); Bh-43, *B.*
348 *humptydooensis* sp. nov. (MSMB43^T); Bh-121. *B. humptydooensis* sp. nov. (MSMB121); Bh-
349 122, *B. humptydooensis sp*. nov. (MSMB122). +, Positive reaction; -, negative reaction. All
350 strains were positive for nitrate, gelatin, glucose, mannose, mannitol, N-acetylglucosamine,
351 gluconate, caprate, adipate, malate, citrate, and phenylacetate. All strains were negative for
352 glucose (acidification) and urea.

Characteristic	Bp*	Bt	Bh-43	Bh-122	Bh-121
Tryptophan	+	-	-	-	-
Arginine	+	-	+	-	-
Aesculin	-	+	+	+	+
PNPG	-	-	-	-	+
Assimilation of:					
Arabinose	-	+	+	+	+
Maltose	-	+	-	-	-

353 * Data obtained from Wuthiekanun et al., (1996).

354

Table 2. Summary of MIC's determined in triplicate by broth microdilution method.

Concentrations are given in mg/L.

Antimicrobial substance	*B. pseudomallei* MIC_{90}	MSMB43[T]	MSMB121	MSMB122
Amoxicillin/clavulanic acid	8[a]	32/16[*]	32/16[*]	32/16[*]
Ceftazidime	4[a]	4	2	4
Imipenem	1[a]	0.5	0.5	0.5
Rifampicin	32[a]	>8	>8	8
Chloramphenicol	16[a]	8	4	4
Trimethoprim/sulfamethoxazole	1.5[c]	0.5/9.5	≤0.25/4.75	≤0.25/4.75
Streptomycin	n/a	>32	>32	>32
Gentamicin	64[a]	32	32	>32
Doxycycline	1[a]	1	1	1
Tigecycline	4[b]	2	4	4
Ciprofloxacin	4[a]	0.5	1	1
Levofloxacin	4[a]	0.5	1	1

[*]Resistance was observed based upon the CLSI breakpoints of *B. pseudomallei*.

Data obtained from [a] Thibault *et al.*, (2004) [b] Ahmad *et al*., (2013) [c] Crowe *et al*., (2014).

361   **Table 3.** Whole genome data for *B. pseudomallei* group.

Strain/species	GC content	Genome size	# CDS
MSMB43^T *B. humptydooensis*	67.1%	7.3* Mb	6324
MSMB121 *B. humptydooensis*	67.5%	6.7 Mb	5795
MSMB122 *B. humptydooensis*	67.5%	6.8 Mb	5845
E264 *B. thailandensis*	67.6%	6.7 Mb	5652
C6786 *B. oklahomensis*	66.9%	7.1 Mb	6097
K96243 *B. pseudomallei*	68.1%	7.2 Mb	5948
ATCC23344 *B. mallei*	68.5%	5.8 Mb	5006

Two chromosomes present for all genomes.

* One plasmid present.

CDS = coding DNA sequences

362

363

364

365

366

367

368   **Table 4.** Genome to genome distance calculation (GGDC) in the bottom half matrix (with
369   confidence intervals) and average nucleotide identity (ANI) in the top half matrix that
370   demonstrate whole genome sequence similarities of *B. humptydooensis* sp. nov. strains and other
371   species in the *pseudomallei* complex. Values highlighted yellow represent values above the
372   similarity threshold that defines members of the same species.

373

	MSMB43 *B. humptydooensis*	MSMB121 *B. humptydooensis*	MSMB122 *B. humptydooensis*	E264 *B. thailandensis*	C6786 *B. oklahomensis*	K96243 *B. pseudomallei*	ATCC23344 *B. mallei*
MSMB43 *B. humptydooensis*	100	98.81	98.98	93.56	91.80	91.17	93.22
MSMB121 *B. humptydooensis*	93.1 ± 2.3	100	99.84	93.59	91.80	93.39	93.26
MSMB122 *B. humptydooensis*	93.1 ± 2.3	97.3 ± 0.02	100	93.60	91.86	93.23	93.25
E264 *B. thailandensis*	51.1 ± 3.2	51.4 ± 3.2	51.3 ± 3.2	100	91.48	93.01	93.05
C6786 *B. oklahomensis*	43 ± 3.1	43.1 ± 3.1	43.1 ± 3.1	40.7 ± 3.1	100	91.17	91.33
K96243 *B. pseudomallei*	48.6 ± 3.1	48.6 ± 3.2	48.6 ± 3.2	45.6 ± 3.3	39.6 ± 3.1	100	99.04
ATCC23344 *B. mallei*	48.7 ± 3.1	49.1 ± 3.1	49.1 ± 3.1	45.7 ± 3.2	40.3 ± 3.1	92.5 ± 2.8	100

374
375

376

377

References:

AHMAD, N., HASHIM, R. & MOHD NOOR, A. 2013. The In Vitro Antibiotic Susceptibility of Malaysian Isolates of Burkholderia pseudomallei. *Int J Microbiol,* 2013**,** 121845.

BARNES, J. L. & KETHEESAN, N. 2005. Route of infection in melioidosis. *Emerg Infect Dis,* 11**,** 638-9.

BRETT, P. J., DESHAZER, D. & WOODS, D. E. 1997. Characterization of Burkholderia pseudomallei and Burkholderia pseudomallei-like strains. *Epidemiol Infect,* 118**,** 137-48.

BURKHOLDER, W. H. 1950. Sour Skin, a Bacterial Rot of Onion Bulbs. *Phytopathology,* 40**,** 115-117.

BUTLER, D. 2012. INFECTIOUS DISEASE Viral research faces clampdown. *Nature,* 490**,** 456-456.

CHENG, A. C. & CURRIE, B. J. 2005. Melioidosis: Epidemiology, pathophysiology, and management. *Clinical Microbiology Reviews,* 18**,** 383-+.

COENYE, T. & VANDAMME, P. 2003. Diversity and significance of Burkholderia species occupying diverse ecological niches. *Environmental Microbiology,* 5**,** 719-729.

COMPANT, S., NOWAK, J., COENYE, T., CLEMENT, C. & BARKA, E. A. 2008. Diversity and occurrence of Burkholderia spp. in the natural environment. *Fems Microbiology Reviews,* 32**,** 607-626.

CROWE, A., MCMAHON, N., CURRIE, B. J. & BAIRD, R. W. 2014. Current antimicrobial susceptibility of first-episode melioidosis Burkholderia pseudomallei isolates from the Northern Territory, Australia. *Int J Antimicrob Agents,* 44**,** 160-2.

CURRIE, B. J. 2015. Melioidosis: evolving concepts in epidemiology, pathogenesis, and treatment. *Semin Respir Crit Care Med,* 36**,** 111-25.

DELCHER, A. L., SALZBERG, S. L. & PHILLIPPY, A. M. 2003. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics,* Chapter 10**,** Unit 10 3.

GEE, J. E., GLASS, M. B., NOVAK, R. T., GAL, D., MAYO, M. J., STEIGERWALT, A. G., WILKINS, P. P. & CURRIE, B. J. 2008. Recovery of a Burkholderia thailandensis-like isolate from an Australian water source. *BMC Microbiol,* 8**,** 54.

GLASS, M. B., GEE, J. E., STEIGERWALT, A. G., CAVUOTI, D., BARTON, T., HARDY, R. D., GODOY, D., SPRATT, B. G., CLARK, T. A. & WILKINS, P. P. 2006a. Pneumonia and septicemia caused by Burkholderia thailandensis in the United States. *J Clin Microbiol,* 44**,** 4601-4.

GLASS, M. B., STEIGERWALT, A. G., JORDAN, J. G., WILKINS, P. P. & GEE, J. E. 2006b. Burkholderia oklahomensis sp. nov., a Burkholderia pseudomallei-like species formerly known as the Oklahoma strain of Pseudomonas pseudomallei. *Int J Syst Evol Microbiol,* 56**,** 2171-6.

GODOY, D., RANDLE, G., SIMPSON, A. J., AANENSEN, D. M., PITT, T. L., KINOSHITA, R. & SPRATT, B. G. 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, Burkholderia pseudomallei and Burkholderia mallei. *J Clin Microbiol,* 41**,** 2068-79.

GOEBEL, B. M. & STACKEBRANDT, E. 1994. Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. *Appl Environ Microbiol,* 60**,** 1614-21.

GU, J. Y., ZANG, S. G., SHENG, X. F., HE, L. Y., HUANG, Z. & WANG, Q. 2015. Burkholderia susongensis sp. nov., a mineral-weathering bacterium isolated from weathered rock surface. *Int J Syst Evol Microbiol,* 65**,** 1031-7.

HOWDEN, H. F. & STOREY, R. I. 1992. Phylogeny of the Rhyparini and the New Tribe Stereomerini, with Descriptions of New Genera and Species (Coleoptera, Scarabaeidae, Aphodiinae). *Canadian Journal of Zoology-Revue Canadienne De Zoologie,* 70**,** 1810-1823.

KONSTANTINIDIS, K. T. & TIEDJE, J. M. 2005. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 2567-2572.

424 LIU, X. Y., BISWAS, S., BERG, M. G., ANTAPLI, C. M., XIE, F., WANG, Q., TANG, M. C., TANG, G. L., ZHANG,
425   L. X., DREYFUSS, G. & CHENG, Y. Q. 2013. Genomics-Guided Discovery of Thailanstatins A, B, and
426   C As Pre-mRNA Splicing Inhibitors and Antiproliferative Agents from Burkholderia thailandensis
427   MSMB43. *Journal of Natural Products,* 76**,** 685-693.
428 MEIER-KOLTHOFF, J. P., AUCH, A. F., KLENK, H. P. & GOKER, M. 2013. Genome sequence-based species
429   delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics,* 14**,**
430   60.
431 MORICI, L. A., HEANG, J., TATE, T., DIDIER, P. J. & ROY, C. J. 2010. Differential susceptibility of inbred
432   mouse strains to Burkholderia thailandensis aerosol infection. *Microb Pathog,* 48**,** 9-17.
433 NAWROCKI, E. P., KOLBE, D. L. & EDDY, S. R. 2009. Infernal 1.0: inference of RNA alignments.
434   *Bioinformatics,* 25**,** 1335-1337.
435 PARAMESWARAN, U., BAIRD, R. W., WARD, L. M. & CURRIE, B. J. 2012. Melioidosis at Royal Darwin
436   Hospital in the big 2009-2010 wet season: comparison with the preceding 20 years. *Medical
437   Journal of Australia,* 196**,** 345-348.
438 PAYNE, G. W., VANDAMME, P., MORGAN, S. H., LIPUMA, J. J., COENYE, T., WEIGHTMAN, A. J., JONES, T.
439   H. & MAHENTHIRALINGAM, E. 2005. Development of a recA gene-based identification approach
440   for the the entire Burkholderia genus. *Applied and Environmental Microbiology,* 71**,** 3917-3927.
441 RICHTER, M. & ROSSELLO-MORA, R. 2009. Shifting the genomic gold standard for the prokaryotic species
442   definition. *Proceedings of the National Academy of Sciences of the United States of America,*
443   106**,** 19126-31.
444 SAHL, J. W., CAPORASO, J. G., RASKO, D. A. & KEIM, P. 2014. The large-scale blast score ratio (LS-BSR)
445   pipeline: a method to rapidly compare genetic content between bacterial genomes. *PeerJ*
446   2:e332.
447 SAWANA, A., ADEOLU, M. & GUPTA, R. S. 2014. Molecular signatures and phylogenomic analysis of the
448   genus Burkholderia: proposal for division of this genus into the emended genus Burkholderia
449   containing pathogenic organisms and a new genus Paraburkholderia gen. nov. harboring
450   environmental species. *Front Genet,* 5**,** 429.
451 STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
452   phylogenies. *Bioinformatics*.
453 STONE, J. K., MAYO, M., GRASSO, S. A., GINTHER, J. L., WARRINGTON, S. D., ALLENDER, C. J., DOYLE, A.,
454   GEORGIA, S., KAESTLI, M., BROOMALL, S. M., KARAVIS, M. A., INSALACO, J. M., HUBBARD, K. S.,
455   MCNEW, L. A., GIBBONS, H. S., CURRIE, B. J., KEIM, P. & TUANYOK, A. 2012. Detection of
456   Burkholderia pseudomallei O-antigen serotypes in near-neighbor species. *BMC Microbiol,* 12**,**
457   250.
458 TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: Molecular
459   Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol,* 30**,** 2725-9.
460 THIBAULT, F. M., HERNANDEZ, E., VIDAL, D. R., GIRARDET, M. & CAVALLO, J. D. 2004. Antibiotic
461   susceptibility of 65 isolates of Burkholderia pseudomallei and Burkholderia mallei to 35
462   antimicrobial agents. *J Antimicrob Chemother,* 54**,** 1134-8.
463 TINDALL, B. J., ROSSELLO-MORA, R., BUSSE, H. J., LUDWIG, W. & KAMPFER, P. 2010. Notes on the
464   characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol,* 60**,** 249-
465   66.
466 TRUNCK, L. A., PROPST, K. L., WUTHIEKANUN, V., TUANYOK, A., BECKSTROM-STERNBERG, S. M.,
467   BECKSTROM-STERNBERG, J. S., PEACOCK, S. J., KEIM, P., DOW, S. W. & SCHWEIZER, H. P. 2009.
468   Molecular basis of rare aminoglycoside susceptibility and pathogenesis of Burkholderia
469   pseudomallei clinical isolates from Thailand. *PLoS Negl Trop Dis,* 3**,** e519.

470 VANLAERE, E., BALDWIN, A., GEVERS, D., HENRY, D., DE BRANDT, E., LIPUMA, J. J.,
471   MAHENTHIRALINGAM, E., SPEERT, D. P., DOWSON, C. & VANDAMME, P. 2009. Taxon K, a
472   complex within the Burkholderia cepacia complex, comprises at least two novel species,
473   Burkholderia contaminans sp nov and Burkholderia lata sp nov. *International Journal of*
474   *Systematic and Evolutionary Microbiology,* 59**,** 102-111.
475 VANLAERE, E., LIPUMA, J. J., BALDWIN, A., HENRY, D., DE BRANDT, E., MAHENTHIRALINGAM, E., SPEERT,
476   D., DOWSON, C. & VANDAMME, P. 2008. Burkholderia latens sp. nov., Burkholderia diffusa sp.
477   nov., Burkholderia arboris sp. nov., Burkholderia seminalis sp. nov. and Burkholderia metallica
478   sp. nov., novel species within the Burkholderia cepacia complex. *Int J Syst Evol Microbiol,* 58**,**
479   1580-90.
480 VIAL, L., CHAPALAIN, A., GROLEAU, M. C. & DEZIEL, E. 2011. The various lifestyles of the Burkholderia
481   cepacia complex species: a tribute to adaptation. *Environmental Microbiology,* 13**,** 1-12.
482 WEST, T. E., FREVERT, C. W., LIGGITT, H. D. & SKERRETT, S. J. 2008. Inhalation of Burkholderia
483   thailandensis results in lethal necrotizing pneumonia in mice: a surrogate model for pneumonic
484   melioidosis. *Trans R Soc Trop Med Hyg,* 102 Suppl 1**,** S119-26.
485 WIERSINGA, W. J., DE VOS, A. F., DE BEER, R., WIELAND, C. W., ROELOFS, J. J., WOODS, D. E. & VAN DER
486   POLL, T. 2008. Inflammation patterns induced by different Burkholderia species in mice. *Cell*
487   *Microbiol,* 10**,** 81-7.
488 WUTHIEKANUN, V., SMITH, M. D., DANCE, D. A., WALSH, A. L., PITT, T. L. & WHITE, N. J. 1996.
489   Biochemical characteristics of clinical and environmental isolates of Burkholderia pseudomallei.
490   *J Med Microbiol,* 45**,** 408-12.
491 YABUUCHI, E., KAWAMURA, Y., EZAKI, T., IKEDO, M., DEJSIRILERT, S., FUJIWARA, N., NAKA, T. &
492   KOBAYASHI, K. 2000. Burkholderia uboniae sp. nov., L-arabinose-assimilating but different from
493   Burkholderia thailandensis and Burkholderia vietnamiensis. *Microbiology and Immunology,* 44**,**
494   307-317.

495

MALDI-TOF:

Preparations of bacterial isolates (MSMB43, MSMB121, MSMB122) for matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) were done according to the ethanol/formic acid extraction protocol recommended by the manufacturer (Bruker Daltonics) and as previously described (Mellmann et al., 2008) Briefly, a loopful of bacterial material was evenly dissolved in 300 µl analytical grade water, and 900 µl pure ethanol was added. The cell suspension was centrifuged at $13,000 \times g$ for 2 minutes, and the supernatant was discarded. The centrifugation was repeated, and the residual ethanol was discarded. The pellet was air dried and thoroughly resuspended in 5-50 µl 70% formic acid depending on the size, and, finally, an equal volume of acetonitrile was added. After centrifugation at $13,000 \times g$ for 2 minutes, 1 µl of the supernatant was transferred to the MALDI target plate. After air-drying at room temperature 1 µl of matrix solution (saturated solution of α-cyano-hydroxy-cinnaminic acid in 50% aqueous acetonitrile containing 2.5% trifluoroacetic acid) per spot were was applied. MALDI-TOF MS was conducted using a Microflex LT mass spectrometer (Bruker Daltonics) equipped with a nitrogen laser. All spectra were recorded in linear, positive ion mode across a mass/charge ratio (m/z) of 2,000 to 20,000. The acceleration voltage was 20 kV. Spectra were collected as a sum of 240 shots across a spot. Main spectra were calculated from 8 spectra per strain and used for construction of a score oriented dendrogram by using the BioTyper software (version 2.3, Bruker Daltonics). In these analyses *B. humptydooensis* sp. nov. strains grouped together with *B. ubonensis* from the Bcc and other members of the *B. pseudomallei* complex (Fig. S1).

Fatty acid methyl ester (FAME) profile analysis:

Fatty acid methyl esters analysis was unable to distinguish among the fatty acid profiles from the three *B. humptydooensis* sp. nov. strains and its closely related species (five *B. thailandensis*, two *B. oklahomensis*, and three *B. ubonensis* strains). The fatty acid methyl esters were performed using the MIDI Sherlock® Microbial Identification System by Microbial ID, Inc. (Delaware, USA). Similarity index analysis was used to display the relatedness of FAME profiles produced by these species. The analysis revealed that FAME profiles of the four *B. humptydooensis* sp. nov. strains grouped into two sub-groups, along with strains from other species (Fig. S2). This finding suggests that *B. humptydooensis* sp. nov. had similar FAME profiles with other tested genetically related *Burkholderia* species and this technique cannot be used to distinguish *B. humptydooensis* sp. nov. or distinguish the other closely related species.

In conclusion, we have utilized comprehensive genotyping techniques including 16S rRNA, *recA*, MLST, and whole genome based GGDC to further support the existence of a new species that is distinct but genetically related to the four members of the pseudomallei complex (*B. pseudomallei*, *B. mallei*, *B. thailandensis*, and *B. oklahomensis*). These analyses confirm the speciation of *B. humptydooensis* sp. nov., a soil bacterial saprophyte found in the Northern Territory of Australia where melioidosis is highly endemic. The addition of *B. humptydooensis* sp. nov. as a new member of the pseudomallei complex will benefit evolutionary studies of *B. pseudomallei*, the serious bacterial pathogen that shares a similar ecological niche with this new species.

Fig. S1. Dendrogram demonstrating strain relatedness revealed by MALDI – TOF analysis.
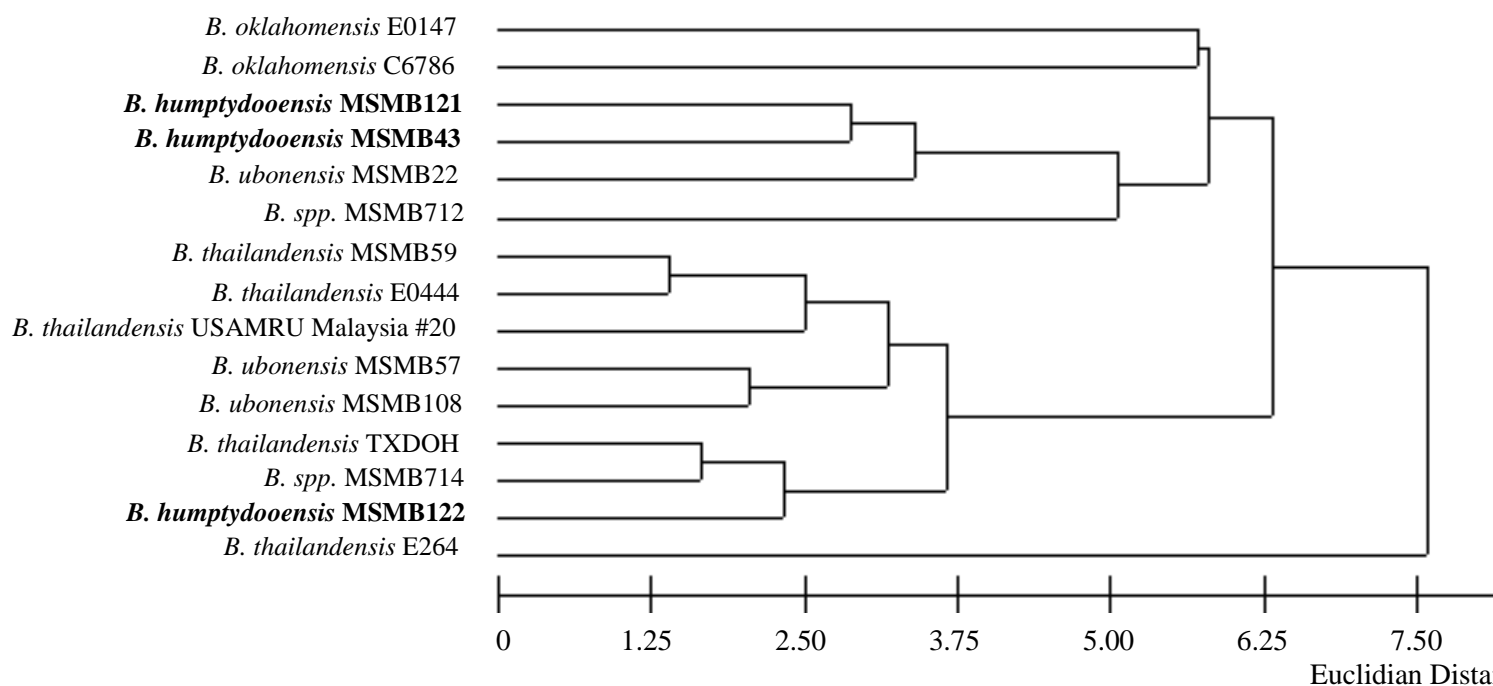
Fig. S2. A dendrogram demonstrating the relatedness of fatty acid compositions in *B. humptydooensis* sp. nov. and other closely related species. Similarity Index in the Microbial Identification System (MIS) was used to display how closely the fatty acid compositions found in *B. humptydooensis* sp. nov., a new species, compared with those from other genetically related species. FAME (Fatty Acid Methyl Esters) profiles were generated using GC-MS. The dendrogram was generated based upon the clustering analysis technique to produce unweighted pair matching based on FAME profiles. Our analysis has shown that FAME profiles from three tested *B. humptydooensis* sp. nov. strains were grouped into two sub-groups. The first group also contained one *B. ubonensis* and two *B. oklahomensis* strains, while the second group contained most of *B. thailandensis* strains and two *B. ubonensis* strains. We noted that *B. thailandensis* E264 had a distinct FAME profile to both subgroups.
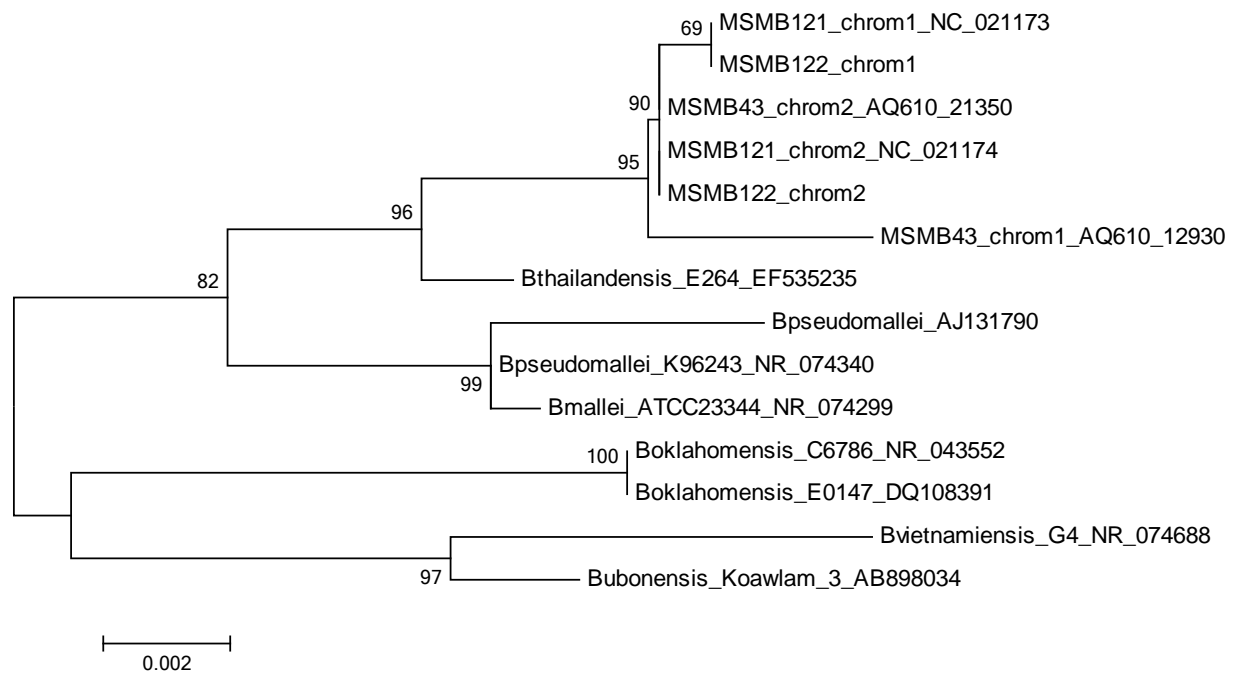
Fig. S3. 16S maximum likelihood (500 bootstrap) phylogeny using both copies of 16S from each *B. humptydooensis* sp. nov. strain (1420 bp and 14 sequences).

Fig. S4. Maximum likelihood phylogeny (1,500 bootstrap) of recA sequence (335 bp) using 193 sequences.